

DERIVATION OF QUANTUM PROBABILITY FROM MEASUREMENT

FEDOR HERBUT*

Serbian Academy of Sciences and Arts, Knez Mihajlova 35, 11000
Belgrade, Serbia

To begin with, it is pointed out that the form of the quantum probability formula originates in the very initial state of the object system as seen when the state is expanded with the eigen-projectors of the measured observable. Making use of the probability reproducibility condition, which is a key concept in unitary measurement theory, one obtains the relevant coherent distribution of the complete-measurement results in the final unitary-measurement state in agreement with the mentioned probability formula. Treating the transition from the final unitary, or premeasurement, state, where all possible results are present, to one complete-measurement result sketchily in the usual way, the well-known probability formula is derived. In conclusion it is pointed out that the entire argument is only formal unless one makes it physical assuming that the quantum probability law is valid in the extreme case of probability-one (certain) events (projectors).

1 Introduction

Probability has no physical meaning if measurement is not taken into account. Hence, the physically most appropriate way to derive probability is to do it in the framework of measurement theory. I have demonstrated advantages of such a procedure within Zurek's way to derive probability from 'envariance' (invariance due to entanglement) [1].

Complete measurement that will be utilized for our derivation consists of two parts: Relevant parts of unitary measurement theory (also called pre-measurement theory or measurement theory short of collapse) and a sketchy phenomenological idea of collapse.

Unitary measurement theory will be along the lines of former work [2], which allowed for redundant entanglement. The basic concepts of this approach, which was based on an unpublished but detailed and systematic

*e-mail: fedorh@sanu.ac.rs

review [3], will be outlined now.

The observables treated in this article are confined to discrete ones, i. e., to ordinary (as opposed to generalized) observables that do not have a continuous part in their spectrum. The object of measurement is denoted by A, and the measuring instrument by B. The measured observable O is given in its unique spectral form (in which, by definition, there is no repetition in the eigenvalues $\{o_k : \forall k\}$:

$$O_A = \sum_k o_k E_A^k. \quad (1a)$$

The eigen-projectors $\{E_A^k : \forall k\}$ satisfy the completeness relation

$$\sum_k E_A^k = I_A, \quad (1b)$$

where I_A is the identity operator in the state space of the object subsystem.

The measuring instrument has a suitable initial state $|\phi\rangle_B^i$ and a so-called pointer observable, which in its unique spectral form reads

$$P_B = \sum_k p_k F_B^k. \quad (2a)$$

There is also the completeness relation

$$\sum_k F_B^k = I_B. \quad (2b)$$

The coindexing is due to a one-to-one relation between the spectral form of the measured observable and that of the measuring instrument with the physical meaning that the result o_k (or equivalently the occurrence of E_A^k) is noted by the measuring instrument by the occurrence of the so-called pointer position F_B^k .

Finally, there is the unitary operator U_{AB} that includes the object-measuring-instrument interaction and transforms the initial state $|\phi\rangle_A^i |\phi\rangle_B^i$ of object+measuring instrument into the final state

$$|\Phi\rangle_{AB}^f \equiv U_{AB} \left(|\phi\rangle_A^i |\phi\rangle_B^i \right). \quad (3)$$

Exact measurement (as opposed to approximate measurement) in its general form (as opposed to the particular case of nondemolition measurement or the even more special case of ideal measurement, cf [3]) is defined by the **calibration condition** for discrete observables (cf [4]):

$$\langle \phi |_A^i E_A^{\bar{k}} | \phi \rangle^i = 1 \quad \Rightarrow \quad \langle \Phi |_{AB}^f F_B^{\bar{k}} | \Phi \rangle_{AB}^f = 1, \quad (4)$$

, which can be equivalently rewritten in the more practical form:

$$E_A^{\bar{k}} | \phi \rangle_A^i = | \phi \rangle_A^i \quad | \Rightarrow \quad F_B^{\bar{k}} | \Phi \rangle_{AB}^f = | \Phi \rangle_{AB}^f. \quad (5)$$

(The equivalence of (4) and (5) is easily proved.).

It was shown in previous work [2] that the calibration condition is equivalent to the **dynamical condition**:

$$\forall k : \quad F_B^K U_{AB} \left(| \phi \rangle_A^i | \phi \rangle_B^i \right) = U_{AB} E_A^K \left(| \phi \rangle_A^i | \phi \rangle_B^i \right). \quad (6)$$

(For the reader's convenience the exposition in this article self-contained. To this purpose, the proof of the claimed equivalence is reproduced in Appendix B.)

2 Role of the Probability Reproducibility Condition

For our purpose, let it be pointed out that an arbitrary state $| \phi \rangle_A^i$ of the object has, on account of the completeness relation (1b), the following decomposition:

$$\forall | \phi \rangle_A^i : \quad | \phi \rangle_A^i = \sum_k \| E_A^k | \phi \rangle_A^i \| \times (E_A^k | \phi \rangle_A^i / \| E_A^k | \phi \rangle_A^i \|) \quad (7)$$

where it is understood that if the first factor in a term is zero, that the entire term is zero though the second factor does not make sense.

Further, due to idempotency of the projectors $\{ E_A^k : \forall k \}$,

$$\forall | \phi \rangle_A^i : \quad \| E_A^k | \phi \rangle_A^i \| = (\langle \phi |_A^i E_A^k | \phi \rangle_A^i)^{1/2}. \quad (8)$$

In our derivation this is, excuse the pun, where the Born rule is borne.

Incidentally, the strict form of the Born rule, the most used expression for pure states and the trace rule are all mutually equivalent forms of the

probability law in quantum mechanics (as proved in Appendix A). We are going to derive it.

A key role is played in unitary measurement theory by the so-called **probability reproducibility condition**:

$$\forall |\phi\rangle_A^i : \quad \langle \Phi |_{AB}^f F_B^k | \Phi \rangle_{AB}^f = \langle \phi |_A^i E_A^k | \phi \rangle_A^i. \quad (9)$$

It was shown in previous work [2] how the probability reproducibility condition follows from the calibration condition. (The proof is reproduced in Appendix C.)

Now we can derive the relevant decomposition of the final state. Making use of the completeness relation (2b) and the idempotency of the projectors $\{F_B^k : \forall k\}$, one can write:

$$\begin{aligned} |\Phi\rangle_{AB}^f &= \sum_k F_B^k |\Phi\rangle_{AB}^f = \sum_k \|F_B^k | \Phi \rangle_{AB}^f\| \times F_B^k | \Phi \rangle_{AB}^f / \|F_B^k | \Phi \rangle_{AB}^f\| = \\ &= \sum_k \left(\langle \Phi |_{AB}^f F_B^k | \Phi \rangle_{AB}^f \right)^{1/2} \times F_B^k | \Phi \rangle_{AB}^f / \|F_B^k | \Phi \rangle_{AB}^f\|. \end{aligned}$$

Finally, the probability reproducibility condition (9) gives

$$|\Phi\rangle_{AB}^f = \sum_k \left(\langle \phi |_A^i E_A^k | \phi \rangle_A^i \right)^{1/2} \times F_B^k | \Phi \rangle_{AB}^f / \|F_B^k | \Phi \rangle_{AB}^f\|. \quad (10)$$

3 The Final Steps

In the final steps we have to leave the unitary final state $|\Phi\rangle_{AB}^f$ (cf (3)) and reach the result of **complete measurement** to which corresponds one value of k for an individual object - the so-called collapse of the unitary final state. Unitary quantum mechanics cannot do this (unless we accept the many-worlds interpretation, which we will not do now).

Peres in his book [5] (the last chapter there) speaks of dequantization when it comes to complete measurement. Accepting the Copenhagen interpretation of quantum mechanics, his dequantization consists in the assumption that the pointer-position projectors $\{F_B^k : \forall k\}$ represent classical events. Viewing the completeness relation (2b) classically only one of

the mutually excluding events can happen. Thus the complete measurement results are obtained.

Bell criticized collapse [6] viewing it entirely within quantum mechanics. The quantum entity that has to collapse, written as a density operator is:

$$|\Phi\rangle_{AB}^f \langle\Phi|_{AB}^f = \sum_k \sum_{k'} \left(\langle\phi|_A^i E_A^k |\phi\rangle_A^i \right)^{1/2} \left(\langle\phi|_A^i E_A^{k'} |\phi\rangle_A^i \right)^{1/2} \times \\ \left(F_B^k |\Phi\rangle_{AB}^f / \|F_B^k |\Phi\rangle_{AB}^f\| \right) \left(\langle\Phi|_{AB}^f F_B^{k'} / \|F_B^{k'} |\Phi\rangle_{AB}^f\| \right). \quad (11)$$

One should note that in (11), besides the diagonal ($k=k'$) terms also the off-diagonal ($k \neq k'$) terms are non-zero (each for some initial state). The latter express **coherence**. They must be deleted in collapse. Thus, the first step is replacing the LHS(11) by

$$\rho_{AB} \equiv \sum_k \langle\phi|_A^i E_A^k |\phi\rangle_A^i \times \\ F_B^k |\Phi\rangle_{AB}^f / \|F_B^k |\Phi\rangle_{AB}^f\| \langle\Phi|_{AB}^f F_B^k / \|F_B^k |\Phi\rangle_{AB}^f\|. \quad (12)$$

Bell called ρ_{AB} the "butchered state".

In spite of butchering the coherence in $|\Phi\rangle_{AB}^f \langle\Phi|_{AB}^f$ one would expect that ρ_{AB} still represents the state of individual quantum systems as the former state did. But, in the second step of collapse, one assumes that ρ_{AB} given by (12) describes the state of an ensemble in which the states of the individual systems are described by the pure states in the terms in (12). So that (12) is assumed to represent a **mixture** with the **statistical weights**

$$\forall k : \quad w_k \equiv \langle\phi|_A^i E_A^k |\phi\rangle_A^i. \quad (13)$$

Bell called this step "replacing "or" by "and"".

The final and for our purpose the most important step is assuming that the **probability** of obtaining the result $F_B^k |\Phi\rangle_{AB}^f / \|F_B^k |\Phi\rangle_{AB}^f\|$ in complete measurement **equals the statistical weight** w_k given by (13). This ends our argument of deriving the quantum probability law from general measurement, at least its formal part. It is physically completed in concluding remark C in section 5.

One should note that the steps that we have made use of in this section are actually phenomenological, i. e., we know from experience that these steps are made in complete measurement.

4 The Mixed Initial State Case

Now we assume that the initial state of the object system is a general state. (Our interest lies, of course, in mixed states because we have already dealt with the pure states.) The method called **purification** will be applied to reduce general states to pure states.

We denote the object system by A_1 . Let A_2 be another system, which will play only a formal role.

Let $\rho_{A_1}^i = \sum_i r_i |i\rangle_{A_1} \langle i|_{A_1}$ be a decomposition of the given initial state of the object system A_1 into its positive-eigenvalue norm-one eigenvectors. Further, let $\{|i\rangle_{A_2} : \forall i\}$ be an arbitrary orthonormal set of vectors in the state space of A_2 . We define

$$\forall \rho_{A_1}^i : |\phi\rangle_{A_1 A_2}^i \equiv \sum_i r_i |i\rangle_{A_1} |i\rangle_{A_2}. \quad (14a)$$

The essential property of this composite-system pure state, which characterizes purification, is that

$$\text{tr}_{A_2} \left(|\phi\rangle_{A_1 A_2}^i \langle \phi|_{A_1 A_2}^i \right) = \rho_{A_1}^i, \quad (14b)$$

$\rho_{A_1}^i$ being the initial state of the object subsystem that we started with.

Let the measured observable be $O_{A_1} = \sum_k o_k E_{A_1}^k$, and let the measuring instrument be subsystem B with the initial state $|\phi\rangle_B^i$ and the pointer observable $P_B = \sum_k p_k F_B^k$ as before. Then, as proved in the preceding sections, the probability to obtain in complete measurement of O_{A_1} the state $F_B^k |\Phi\rangle_{A_1 A_2 B} / \|F_B^k |\Phi\rangle_{A_1 A_2 B}\|$ is:

$$\langle \phi|_{A_1 A_2}^i E_{A_1}^k |\phi\rangle_{A_1 A_2}^i. \quad (15)$$

Now it is time for depurification, i. e., to rid ourselves of the passive subsystem A_2 . The expectation value (15) is standardly rewritten in terms of its subsystem state operator (reduced density operator) as:

$$\text{tr} \left(\rho_{A_1}^i E_{A_1}^k \right). \quad (16)$$

This is the final result.

5 Concluding Remarks

A) We have seen in relation (10) that the final state of unitary measurement theory is a state in which all possible result are contained. In order to reach the final state of complete measurement the two steps described sketchily in section 3 are unavoidable: one must terminate the coherence in (10) (the "butchering" following Bell), and then the drastic change that the butchered state ρ_{AB} is not valid for individual systems, only for an ensemble of such, where the terms in (12) apply to the individual systems making up the ensemble (Bell's "replacing "and" by "or"").

The fact that derivation of a final state of complete measurement is considered to be impossible in unitary quantum mechanics is known as the paradox of quantum measurement. (Though the derivation is possible in the many-worlds interpretation of quantum mechanics, which is not universally accepted.)

B) One might think of complete measurement that does not end in the state $F_B^k |\Phi\rangle_{AB}^f / \|F_B^k |\Phi\rangle_{AB}^f\|$. This may be the case, e. g., if one has overmeasurement [7]. But then one deals with a different probability formula. The one derived in this study, which is the standard one (cf Appendix A) is better understood by the following explanation.

Utilizing (3), one can rewrite the dynamical condition (6) as follows

$$\forall |\phi\rangle_A^i, \forall k : |\phi\rangle_A^i \rightarrow U_{AB}(E_A^k |\phi\rangle_A^i |\phi\rangle_B^i) = F_B^k |\Phi\rangle_{AB}^f. \quad (17)$$

One can see that **each initial term** $E_A^k |\phi\rangle_A^i$ (cf relation (7)) **evolves** (applying to it $U_{AB}(\dots \otimes |\phi\rangle_B^i)$) **separately**, i. e., independently of the rest of the terms, into the corresponding final term $F_B^k |\Phi\rangle_{AB}^f$. Thus, in unitary measurement we have a set of **complete-measurement branches, each evolving independently of each other, but tied up into a whole by coherence**.

Thus, we actually consider one **entire** branch in seeking to reach the corresponding complete-measurement state. We begin with a definite eigenvalue state $E_A^k |\phi\rangle_A^i / \|E_A^k |\phi\rangle_A^i\|$ with $\|E_A^k |\phi\rangle_A^i\| = (\langle \phi |_A^i E_A^k |\phi\rangle_A^i)^{1/2}$ (where the Born rule begins, as stated - cf relation (8) and beneath it).

In overmeasurement we would not start with the entire branch k . One would have $E_A^k = \sum_{\bar{k}} E_{\bar{k}}^k$ and one would endeavor to reach the complete-measurement state corresponding to a fixed \bar{k} value. In the end, one would

then derive $\langle \phi |_A^i \bar{E}_A^k | \phi \rangle_A^i$.

C) The entire derivation in sections 2 and 3 is algebraic and formal. We must put in a suitable physical assumption at the beginning, so that we obtain a physically meaningful result at the end.

Since we have made essential use of the dynamical condition (6), and it is equivalent to the calibration condition (5), it is the latter that must be given physical meaning. To do this the idea of a (statistically) sharp value must be expressible as $\langle \phi |_A^i E_A^k | \phi \rangle_A^i = 1$. Since the latter is equivalent to $E_A^k | \phi \rangle_A^i = 1 \times | \phi \rangle_A^i$, we must **assume** that if an event (projector) in a pure state has the eigenvalue one, then the event is **certain in this state**.

Thus, assuming the physical validity of the probability formula that is to be derived in the special extreme case, we obtain the physically meaningful final formula for all cases (13).

I have read somewhere that you cannot derive probability unless you put in something of probability. It is certainly valid for our derivation. Incidentally, a completely different derivation of the quantum probability law [8] started with the same physical assumption.

Appendix A. Equivalent forms of the quantum probability law

Let P denote a projector and let $|\psi\rangle$ and $|\phi\rangle$ denote norm-one vectors. The following three probability expressions are equivalent:

$$\langle \psi | P | \psi \rangle \quad (1) \quad \Leftrightarrow \quad |\langle \psi | \phi \rangle|^2 \quad (2) \quad \Leftrightarrow \quad \text{tr}(P | \psi \rangle \langle \psi |) \quad (3). \quad (A.1)$$

Expression (2) is the "Born rule" (in the strict sense), and expression (3) is called the "trace rule".

Proof. We assume that $P = | \phi \rangle \langle \phi |$. Then expression (1) becomes expression (2) as one can see using the Dirac rules.

Let $P = \sum_k | \phi, k \rangle \langle \phi, k |$ be a complete orthogonal decomposition of P . Let us further assume that the probability of an orthogonal sum (disjoint events) is sum of the probabilities of the terms. Then

$$\sum_k |\langle \psi | \phi, k \rangle|^2 = \langle \psi | \left(\sum_k | \phi, k \rangle \langle \phi, k | \right) | \psi \rangle = \langle \psi | P | \psi \rangle. \quad (A.2)$$

The first equivalence is proved.

Having in mind evaluation of the trace in a basis in which $|\psi\rangle$ is one of the basis vectors, one can see that

$$\langle \psi | P | \psi \rangle = \text{tr}(P | \psi \rangle \langle \psi |). \quad (A.3)$$

This proves the equivalence of (1) with (3). The second equivalence in (A.1) is then a consequence of transitivity of equivalences. \square

Appendix B. Proof of the dynamical condition

We now express and prove the **dynamical condition**, valid for general measurement, and being **equivalent to the calibration condition**.

The **claim** goes as follows.

One has exact measurement **if and only if**

$$\forall |\phi\rangle_A^{\mathbf{i}}, \forall k : \left(F_B^k U_{AB} \right) \left(|\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} \right) = \left(U_{AB} E_A^k \right) \left(|\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} \right) \quad (B.1)$$

is valid.

One *proves necessity* as follows. The completeness relation $\sum_{k'} E_A^{k'} = I_A$, use of the calibration condition (5), and orthogonality and idempotency of the F_B^k projectors enable one to write for each k value :

$$\begin{aligned} F_B^k U_{AB} |\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} &= \\ \sum_{k'} \|E_A^{k'} |\phi\rangle_A^{\mathbf{i}}\| \times F_B^k U_{AB} \left(E_A^{k'} |\phi\rangle_A^{\mathbf{i}} / \|E_A^{k'} |\phi\rangle_A^{\mathbf{i}}\| \right) |\phi\rangle_B^{\mathbf{i}} &= \\ \sum_{k'} \|E_A^{k'} |\phi\rangle_A^{\mathbf{i}}\| \times F_B^k \mathbf{F}_B^{k'} U_{AB} \left(E_A^{k'} |\phi\rangle_A^{\mathbf{i}} / \|E_A^{k'} |\phi\rangle_A^{\mathbf{i}}\| \right) |\phi\rangle_B^{\mathbf{i}} &= \\ \|E_A^k |\phi\rangle_A^{\mathbf{i}}\| \times F_B^k U_{AB} \left(E_A^k |\phi\rangle_A^{\mathbf{i}} / \|E_A^k |\phi\rangle_A^{\mathbf{i}}\| \right) |\phi\rangle_B^{\mathbf{i}}. \end{aligned}$$

Thus,

$$F_B^k U_{AB} |\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} = \|E_A^k |\phi\rangle_A^{\mathbf{i}}\| \times F_B^k U_{AB} \left(E_A^k |\phi\rangle_A^{\mathbf{i}} / \|E_A^k |\phi\rangle_A^{\mathbf{i}}\| \right) |\phi\rangle_B^{\mathbf{i}}. \quad (B.2)$$

Finally, on account of (5) again, we can omit F_B^k , so that, after cancellation, one obtains:

$$F_B^k U_{AB} |\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} = U_{AB} E_A^k |\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}}.$$

The cancellation cannot be done if $\|E_A^k |\phi\rangle_A^{\mathbf{i}}\| = 0$. But the claimed relation (B.1) is still valid because the RHS is obviously zero, and so is the LHS as seen in (B.2).

To *prove sufficiency*, let

$$\left(U_{AB} E_A^k \right) \left(|\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} \right) = \left(F_B^k U_{AB} \right) \left(|\phi\rangle_A^{\mathbf{i}} |\phi\rangle_B^{\mathbf{i}} \right)$$

be valid for all k values, and let $|\phi\rangle_A^i = E_A^{\bar{k}} |\phi\rangle_A^i$ be satisfied for a fixed value $k \equiv \bar{k}$. Then, one has in particular

$$(U_{AB} E_A^{\bar{k}}) (|\phi\rangle_A^i |\phi\rangle_B^i) = (F_B^{\bar{k}} U_{AB}) (|\phi\rangle_A^i |\phi\rangle_B^i).$$

One can here omit $E_A^{\bar{k}}$ due to the assumed definite value in $|\phi\rangle_A^i$ (cf (5)), and thus the explicit form of the calibration condition (5) is obtained. *This ends the proof.*

Appendix C. Proof of the Probability Reproducibility Condition

The probability reproducibility condition reads:

$$\forall |\phi\rangle_A^i, \forall k: \langle \Phi |_{AB}^f F_B^k | \Phi \rangle_{AB}^f = \langle \phi |_A^i E_A^k | \phi \rangle_A^i, \quad (C.1)$$

Proof. Utilizing definition (3), the dynamical condition (6), and the idempotency of F_B^k and of E_A^k , one can see that

$$LHS(C.1) = \langle \phi |_A^i \langle \phi |_B^i (E_A^k U_{AB}^\dagger) (U_{AB} E_A^k) | \phi \rangle_A^i | \phi \rangle_B^i = RHS(13).$$

References

- [1] F. Herbut, *Eur. Phys. J. Plus* **127**, 14 (2012).
- [2] F. Herbut, *Int. J. Quant. Inf.* **12**, 1450032 (16 pages) (2014).
- [3] F. Herbut, arXiv:1412.7862 (2014).
- [4] P. Busch, P. K. Lahti, and P. Mittelstaedt, *The Quantum Theory of Measurement*, 2nd edition (Springer, Berlin, 1996).
- [5] A. Peres, *Quantum Theory: Concepts and Methods*, (Kluwer Ac. Publ., Dordrecht, 1993).
- [6] J. Bell, *Physics World* August, 33 (1990).
- [7] F. Herbut, arXiv:1511.07402 (2015).
- [8] E. Farhi and J. GOLDSTONE, *Ann. Phys. (N. Y.)* **192**, 368(1989).